ADVANCES IN APPLIED MATHEMATICS 7, 170–181 (1986)

Empirical Bayes Stock Market Portfolios

THOMAS M. COVER* AND DAVID H. GLUSS†

Department of Statistics, Stanford University, Stanford, California 94305

DEDICATED TO HERBERT ROBBINS ON HIS 70TH BIRTHDAY

We exhibit a compound sequential Bayes portfolio selection algorithm based solely on the past which not only lives off market fluctuations but follows the drift as well. In fact, this sequential portfolio performs as well (up to first order terms in the exponent) as the optimal portfolio based on advance knowledge of the n-period empirical distribution of the market. Moreover, to first order in the exponent, the capital resulting from this portfolio will be no less than the best of the available stocks. This is a result that holds for every sample sequence. Thus bull markets and bear markets can not fool the investor into over-committing or under-committing his capital to the risky alternatives available to him. The goal is accomplished by a choice of portfolio which is robust with respect to futures that may differ drastically from the past.

1. INTRODUCTION

We consider sequential investments in a stock market with the goal of performing as well as if we knew the empirical distribution of future market performance.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_m) \geq 0$ denote a market vector for one investment period, where x_i is the number of units returned from an investment of 1 unit in the ith stock. A portfolio $\mathbf{b} = (b_1, b_2, \ldots, b_m)$, $b_i \geq 0$, $\Sigma b_i = 1$, is the proportion of the current capital invested in each of the m stocks. Thus $S = \mathbf{b}^t\mathbf{x} = \Sigma b_i x_i$ is the factor by which the capital is increased in one investment period using portfolio $\mathbf{b}$.

For the moment let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be independent identically distributed random vectors drawn according to $F(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^m$, where F is some known

*Stanford University, Department of Statistics and Electrical Engineering. This work was partially supported by NSF Grant ECS82-11568.

†Stanford University. Now at LSI Logic.

0196-8858/86 $7.50

distribution function. Let $S_n = \prod_{i=1}^n \mathbf{b}'\mathbf{X}_i$ denote the capital at time n resulting from an initial capital $S_0 = 1$, and a reinvestment of assets according to portfolio $\mathbf{b}$ at each investment opportunity. Then

$$S_n = \prod_{i=1}^{n} \mathbf{b}'\mathbf{X}_i = e^{\sum_{i=1}^n \ln \mathbf{b}'\mathbf{X}_i}$$

$$= e^{n(E \ln \mathbf{b}'\mathbf{X} + o_p(1))}, \tag{1.1}$$

by the strong law of large numbers, where $o_p(1) \to 0$, a.e. We observe from (1.1) that, to first order in the exponent, the growth rate of capital S_n is determined by the expected log return. Thus motivated, we define

$$W(\mathbf{b}, F) = \int \ln \mathbf{b}'\mathbf{x}\, dF(\mathbf{x}) \tag{1.2}$$

as the expected log return for portfolio $\mathbf{b}$ and stock distribution $F(\mathbf{x})$. Let $\mathbf{b}^*(F)$ maximize $W(\mathbf{b}, F)$ over portfolios $\mathbf{b}$ and let this maximum be denoted by

$$W^*(F) = \max_{\mathbf{b}} W(\mathbf{b}, F)$$

$$= W(\mathbf{b}^*(F), F). \tag{1.3}$$

It follows for $\mathbf{X}_1, \mathbf{X}_2, \ldots,$ i.i.d. $\sim F$ that $\mathbf{b}^*(F)$ achieves an exponential growth rate of capital with exponent $W^*(F)$. Moreover Breiman [1] establishes

$$\overline{\lim}\, \frac{1}{n} \ln S_n \leq W^*(F), \qquad \text{a.e.}, \tag{1.4}$$

for any other portfolio sequence, time-invariant or otherwise. Thus $\mathbf{b}^*(F)$ is asymptotically optimal in this sense, and $W^*(F)$ is the highest possible exponent for the growth rate of capital.

Now suppose that the stock vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are not random. Let $F_k(\mathbf{x})$ denote the empirical cumulative probability distribution function induced by $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Thus F_k corresponds to a uniform distribution with mass $1/k$ on each of the points $\mathbf{x}_i$. If the constant portfolio $\mathbf{b}_k = \mathbf{b}$ is used for each investment period, the resulting capital at time n is given by

$$S_n = \prod_{i=1}^{n} \mathbf{b}'\mathbf{x}_i$$

$$= e^{n(1/n \sum_{i=1}^n \ln \mathbf{b}'\mathbf{x}_i)}$$

$$= e^{nW(\mathbf{b}, F_n)}$$

$$\leq e^{nW(\mathbf{b}^*(F_n), F_n)}$$

$$= e^{nW^*(F_n)}. \tag{1.5}$$

Consequently, no constant portfolio, even with prior knowledge of n and F_n, can achieve capital S_n larger than $e^{nW^*(F_n)}$

We are thus motivated to find a sequential portfolio selection algorithm that ε-achieves $W^*(F_n)$, where $W^*(F_n)$ represents the maximal expected log return over all time-invariant portfolios $\mathbf{b}$. In general, we cannot hope to achieve $W(\mathbf{b}^*(F_n), F_n)$, since F_n is not known during the sequence of n investments.

Here $W^*(F_n)$ plays the role of the Bayes envelope in the compound sequential Bayes theory developed by Robbins [2, 3, 4]. In Robbins' theory we desire estimators performing ε-Bayes with respect to the empirical distribution on the underlying parameters. Here we wish to ε-achieve the Bayes capital growth rate with respect to the empirical distribution of the market. Blackwell's geometric techniques [5, 6] will be used in the proof. Further discussion of the relevance of the log optimal envelope $W^*(F_n)$ may be found in [7–13].

2. Preliminary Properties

As before, let

$$W(\mathbf{b}, F) = \mathop{E}_{F} \ln \mathbf{b}'\mathbf{X} = \int \ln \mathbf{b}'\mathbf{x}\, dF(\mathbf{x}) \tag{2.1}$$

denote the expected log return for portfolio $\mathbf{b}$ and distribution F. We shall use the following properties throughout the paper.

LEMMA 1. *$W(\mathbf{b}, F)$ is concave in $\mathbf{b}$ and linear in F.*

Proof. Linearity is obvious and concavity follows from Jensen's inequality.

LEMMA 2. *If $W(\mathbf{b}^*, F) = \max_b W(\mathbf{b}, F)$, then*

$$\begin{aligned} E\left(X_i/\mathbf{b}^{*\prime}\mathbf{X}\right) &= 1, \qquad \text{for } b_i^* > 0 \\ &\le 1, \qquad \text{for } b_i^* = 0. \end{aligned} \tag{2.2}$$

Proof. These are the Kuhn–Tucker conditions. See Bell–Cover [14, 15], or Finkelstein and Whitley [16].

LEMMA 3. *$W^*(F)$ is convex in F.*

Proof. $W(\mathbf{b}, F)$ is linear in F for each portfolio $\mathbf{b}$. The set of pairs $(F, W^*(F))$ is simply the upper envelope of the set of linear manifolds

$(F, W(\mathbf{b}, F))$ indexed by $\mathbf{b}$, and is therefore convex. More precisely, let

$$F_\lambda = \lambda F_1 + \bar{\lambda} F_2, \qquad 0 \le \lambda \le 1. \tag{2.3}$$

By optimality of $\mathbf{b}^*(F)$ with respect to F,

$$W^*(F_1) \ge W(\mathbf{b}^*(F_\lambda), F_1)$$

and

$$W^*(F_2) \ge W(\mathbf{b}^*(F_\lambda), F_2).$$

Thus, using the linearity of $W(\mathbf{b}, F)$ in F,

$$\begin{aligned}\lambda W^*(F_1) + \bar{\lambda} W^*(F_2) &\ge \lambda W(\mathbf{b}^*(F_\lambda), F_1) + \bar{\lambda} W(\mathbf{b}^*(F_\lambda), F_2)\\ &= W(\mathbf{b}^*(F_\lambda), F_\lambda)\\ &= W^*(F_\lambda),\end{aligned} \tag{2.4}$$

establishing the desired convexity.

Throughout this paper, $\mathbf{b}^*(F), \mathbf{b}^*(\mu), \mathbf{b}^*(\mathbf{p})$ are defined to be the expected log optimal portfolios, as specified in (1.3), with respect to distributions F, measures μ, and probability mass functions $\mathbf{p}$.

LEMMA 4. *If the random stock vector* $\mathbf{X} \in \mathbf{R}^m$ *is bounded* ($|\log X_i| \le L$, $i = 1, 2, \ldots, m$, *a.e.*), *then* $W(\mathbf{b}^*(\mu), \mu)$ *is uniformly continuous in variation norm over probability measures* μ.

Proof. Using the fact that $W(\mathbf{b}, \nu) \le W(\mathbf{b}^*(\nu), \nu)$, for all $\mathbf{b}$, we have

$$\begin{aligned}&W^*(\mu) - W^*(\nu)\\ &\quad = W(\mathbf{b}^*(\mu), \mu) - W(\mathbf{b}^*(\mu), \nu) + W(\mathbf{b}^*(\mu), \nu) - W(\mathbf{b}^*(\nu), \nu)\\ &\quad \le \left| \int \log \mathbf{b}^*(\mu)^t \mathbf{x} (\mu - \nu)(d\mathbf{x}) \right| + W(\mathbf{b}^*(\mu), \nu) - W(\mathbf{b}^*(\nu), \nu)\\ &\quad \le \int \left| \log \mathbf{b}^*(\mu)^t \mathbf{x} \right| |(\mu - \nu)(d\mathbf{x})|\\ &\quad \le L \int |\mu - \nu| = L \|\mu - \nu\|_1.\end{aligned} \tag{2.5}$$

By symmetry, we have

$$W^*(\nu) - W^*(\mu) \le L \|\mu - \nu\|_1. \tag{2.6}$$

Thus

$$|W^*(\mu) - W^*(\nu)| \le L \|\mu - \nu\|_1. \tag{2.7}$$

Finally, we shall specialize Blackwell's approach–exclusion theorem [5, 6] for our needs. The following theorem says roughly that the running average of points successively drawn from the farthest separating hyperplane (between the current average point and a given convex set) converges to that convex set. Note that the farthest separating hyperplane is a support hyperplane for the convex set.

LEMMA 5 (Blackwell). *Let* $\mathbf{A}$ *be a closed convex subset of a bounded region* $\mathbf{U} \subset \mathbf{R}^d$. *Consider a sequence of points* $\{\mathbf{q}_i\}$ *chosen according to the following scheme: Let* $\mathbf{q}_1 = \mathbf{y}_1$ *be any point in the region* $\mathbf{U}$. *For* $n \geq 2$, *and* $\mathbf{q}_n \notin A$, *let* $\mathbf{t}_n$ *denote the closest point in* $\mathbf{A}$ *to* $\mathbf{q}_n$. *(The point* $\mathbf{t}_n$ *is unique because of the convexity of* $\mathbf{A}$.) *Let* H_n *denote the supporting hyperplane to* $\mathbf{A}$ *passing through the point* $\mathbf{t}_n$ *and orthogonal to* $\mathbf{q}_n - \mathbf{t}_n$. *If* $\mathbf{q}_n \in \mathbf{A}$, *let* H_n *be any supporting hyperplane of* $\mathbf{A}$.

Let $\mathbf{y}_{n+1}$ be any point in $\mathbf{U}$ on the hyperplane H_n and let

$$\mathbf{q}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} \mathbf{y}_i. \tag{2.8}$$

Then $\mathbf{q}_n \to \mathbf{A}$. If the diameter of $\mathbf{U}$ is D, then

$$\|\mathbf{q}_n - \mathbf{A}\| \leq D/\sqrt{n}, \qquad \textit{for all } n. \tag{2.9}$$

Proof. Let $d_n = \|\mathbf{q}_n - \mathbf{t}_n\|$, where $\|\cdot\|$ denotes Euclidean distance. Then, if $\mathbf{q}_n \notin \mathbf{A}$,

$$\begin{aligned}
d_{n+1}^2 &= \|\mathbf{q}_{n+1} - \mathbf{t}_{n+1}\|^2 \\
&\leq \|\mathbf{q}_{n+1} - \mathbf{t}_n\|^2 \\
&= \left\| \frac{n}{n+1}\mathbf{q}_n + \frac{1}{n+1}\mathbf{y}_{n+1} - \mathbf{t}_n \right\|^2 \\
&= \left\| \frac{n}{n+1}(\mathbf{q}_n - \mathbf{t}_n) + \frac{1}{n+1}(\mathbf{y}_{n+1} - \mathbf{t}_n) \right\|^2 \\
&= \frac{n^2}{(n+1)^2} d_n^2 + \frac{1}{(n+1)^2} \|\mathbf{y}_{n+1} - \mathbf{t}_n\|^2 \\
&\leq \frac{n^2}{(n+1)^2} d_n^2 + \frac{D^2}{(n+1)^2},
\end{aligned} \tag{2.10}$$

where we have used

$$\mathbf{q}_{n+1} = (n/(n+1))\mathbf{q}_n + (1/(n+1))\mathbf{y}_{n+1}, \tag{2.11}$$

and the orthogonality of $\mathbf{q}_n - \mathbf{t}_n$ and $\mathbf{y}_{n+1} - \mathbf{t}_n$. If $\mathbf{q}_n \notin \mathbf{A}$, (2.10) is true, since $d_n = 0$. Now if $d_n^2 \le D^2/n$ (true for $n = 1$), then

$$\begin{aligned} d_{n+1}^2 &\le \frac{D^2}{(n+1)^2} + \left(\frac{n}{n+1}\right)^2 \frac{D^2}{n} \\ &= D^2/(n+1), \end{aligned} \tag{2.12}$$

and the lemma is proved by induction.

3. Compound Sequential Bayes Portfolios

In this section, the stock vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ will have no underlying distribution. However, we shall constrain the sequence to take values in some finite set $\mathbf{X}$. Our bounds will depend on the cardinality of this set.

Theorem 1. *There exists a sequence of portfolios $\mathbf{b}_k$, where $\mathbf{b}_k$ depends only on the past $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k-1}$ and the cardinality of $\mathbf{X}$, such that the cumulative log return satisfies*

$$\frac{1}{n} \log S_n = \frac{1}{n} \sum_{k=1}^{n} \ln \mathbf{b}_k' \mathbf{x}_k \ge W(\mathbf{b}(F_n), F_n) - \frac{c}{\sqrt{n}}, \tag{3.1}$$

for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbf{X}$ and for all n, where the constant c depends only on the range $\mathbf{X}$.

Remark. Thus one can perform asymptotically as well on sequential investments as if one knew F_n ahead of time. In particular, we can compare this result to what could be achieved by looking at the newspaper n investment periods in the future. The price of the ith stock at time n relative to its price today is given by $\prod_{k=1}^{n} x_{ik}$, the product of the factors by which it increases over the n periods. By comparing stock prices then and now, we could determine the stock that increased the most and invest all of our capital in that stock. If we leave that investment untouched, the resulting capital $\tilde{S}_n$ is given by

$$\begin{aligned} \tilde{S}_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) &= \max_{i=1,2,\ldots,m} x_{i1} x_{i2} \ldots x_{in} \\ &= e^{n(\max_i 1/n \sum_{k=1}^{n} \ln x_{ik})} \\ &= e^{n \max_{b \in B_0} W(\mathbf{b}, F_n)}, \end{aligned} \tag{3.2}$$

where $\mathbf{B}_0$ is the set of portfolios $\mathbf{B}_0 = \{\mathbf{b} \in \mathbf{B}\colon b_i = 1, \text{ some } i\}$. Thus, since

$\mathbf{B}_0 \subseteq \mathbf{B}$, the clairvoyant investor achieves capital

$$\tilde{S}_n(x_1, x_2, \ldots, x_n) \le e^{nW(\mathbf{b}^*(F_n), F_n)}. \tag{3.3}$$

Consequently, $(1/n)\ln \tilde{S}$ is always no greater than and usually substantially less than the asymptotically achievable goal $W(\mathbf{b}^*(F_n), F_n)$. Thus the sequential portfolio algorithm outperforms an investor who has knowledge of the individual stock prices in the far future.

Proof. Let $\mathbf{X} = \{\mathbf{a}_1, \ldots, \mathbf{a}_M\}$ denote the range of stock vectors in $\mathbf{R}^m$. Let $\mathbf{L} = \max_{i=1,\ldots,M,\, j=1,2,\ldots,m} |\log a_{ij}|$. Let $\mathbf{p} = (p(\mathbf{a}_1), p(\mathbf{a}_2), \ldots, p(\mathbf{a}_M))$, $p \ge 0$, $\Sigma p = 1$, denote a probability vector in $\mathbf{R}^M$. Let

$$\mathbf{p}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}(\mathbf{x}_i = \mathbf{x}), \qquad \mathbf{x} \in \mathbf{X}, \tag{3.4}$$

denote the empirical probability mass vector generated by the sequence of stock vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbf{X}$, where $\mathbf{I}(\mathbf{x})$ is the indicator M-vector with a 1 in the xth place and 0's elsewhere.

A sequential choice of portfolios $\mathbf{b}_1, \mathbf{b}_2(\mathbf{x}_1), \ldots, \mathbf{b}_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1})$ results in capital

$$S_n = \prod_{k=1}^{n} \mathbf{b}_k^t(\mathbf{x}_1, \ldots, \mathbf{x}_{k-1})\mathbf{x}_k \tag{3.5}$$

at time n. Let

$$W_n = \frac{1}{n}\ln S_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{b}_k^t \mathbf{x}_k \tag{3.6}$$

denote the exponent of this capital. That is,

$$S_n = e^{nW_n}. \tag{3.7}$$

Here W_n is the cumulative average daily log return associated with $\mathbf{b}_k(\cdot)$. We seek to drive W_n above $W^*(F_n)$ by appropriate choice of $\mathbf{b}_k$.

We now invoke Blackwell's lemma with $d = M + 1$, where M is the cardinality of $\mathbf{X}$. The set

$$\mathbf{U} = \left\{(\mathbf{p}, W) : \mathbf{p} \in \mathbf{R}^M,\ p \ge 0, \sum p(\mathbf{a}) = 1,\ -L \le W \le L\right\} \subset \mathbf{R}^d \tag{3.8}$$

is given by the Cartesian product of the probability M-simplex with the interval $[-L, L]$. The diameter D of $\mathbf{U}$ is

$$D = \sqrt{4L^2 + 2}\,. \tag{3.9}$$

The convex set $\mathbf{A}$ in the lemma is the Bayes envelope

$$\mathbf{A} = \left\{(\mathbf{p}, W): W \geq W^*(\mathbf{p}), \mathbf{p} \in \mathbf{R}^M, p_i \geq 0, \sum p_i = 1\right\}. \tag{3.10}$$

Now consider the situation at time n. Let

$$\mathbf{q}_n = (\mathbf{p}_n, W_n) \in \mathbf{R}^{M+1} \tag{3.11}$$

denote the current empirical probability mass vector on $\mathbf{X}$ and the current empirical average log return $W_n = (1/n)\log S_n$. As in Blackwell's lemma, let $\mathbf{t}_n$ denote the closest point in (the Bayes envelope) $\mathbf{A}$ to $\mathbf{q}_n$. Denote the components of $\mathbf{t}_n$ by

$$\mathbf{t}_n = (\hat{\mathbf{p}}_n, \hat{W}_n). \tag{3.12}$$

By construction, if $\mathbf{q}_n \notin \mathbf{A}$, we will have $\hat{W}_n = W^*(\hat{\mathbf{p}}_n)$, i.e., $\mathbf{t}_n$ is on the boundary of $\mathbf{A}$. This defines $\hat{\mathbf{p}}_n$ and $\mathbf{b}^*(\hat{\mathbf{p}}_n)$ as a function of $\mathbf{q}_n = (\mathbf{p}_n, W_n)$, where $\mathbf{b}^*(\hat{\mathbf{p}}_n)$ is the log optimal portfolio with respect to the probability mass function $\hat{\mathbf{p}}_n$. Note that the portfolio $\mathbf{b}^*(\hat{\mathbf{p}}_n)$ generates the supporting hyperplane

$$H_n = \left\{(\mathbf{p}, W(\mathbf{b}^*(\hat{\mathbf{p}}_n), \mathbf{p})), \mathbf{p} \in \mathbf{R}^M\right\} \tag{3.13}$$

for the Bayes log return envelope $\mathbf{A}$. We now set

$$\mathbf{b}_{n+1} = \mathbf{b}^*(\hat{\mathbf{p}}_n), \tag{3.14}$$

and assert that this portfolio achieves the goal of the theorem.

Defining

$$\mathbf{y}_n = \left(\mathbf{I}(\mathbf{x}_n), \ln \mathbf{b}_n^t \mathbf{x}_n\right), \tag{3.15}$$

we observe that

$$\mathbf{q}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{y}_i, \tag{3.16}$$

and

$$\mathbf{q}_{n+1} = \frac{n}{n+1}\mathbf{q}_n + \frac{1}{n+1}\mathbf{y}_{n+1}, \tag{3.17}$$

where

$$\mathbf{y}_{n+1} \in H_n. \tag{3.18}$$

Thus the conditions of the lemma are satisfied. Consequently,

$$\|\mathbf{q}_n - \mathbf{A}\| \le D/\sqrt{n}\,. \tag{3.19}$$

We now show that the last component $W_n = (1/n)\log S_n$ of $\mathbf{q}_n = (\mathbf{p}_n, W_n)$ converges to the desired value $W^*(\mathbf{p}_n)$. Let $\|\cdot\|_p$ denote the p-norm. From (3.11), (3.12), and (3.19),

$$\|\mathbf{p}_n - \hat{\mathbf{p}}_n\| \le \|(\mathbf{p}_n, W_n) - (\hat{\mathbf{p}}_n, \hat{W}_n)\| = \|\mathbf{q}_n - \mathbf{A}\| \tag{3.20}$$

implies

$$\|\hat{\mathbf{p}}_n - \mathbf{p}_n\|_2 \le D/\sqrt{n}\,, \tag{3.21}$$

which, by the Cauchy–Schwartz inequality, implies

$$\|\hat{\mathbf{p}}_n - \mathbf{p}_n\|_1 \le D\sqrt{M}/\sqrt{n}\,. \tag{3.22}$$

By Lemma 4, it follows that

$$|W^*(\hat{\mathbf{p}}_n) - W^*(\mathbf{p}_n)| \le \frac{LD\sqrt{M}}{\sqrt{n}}. \tag{3.23}$$

By the triangle inequality and (3.23),

$$\begin{aligned} |W_n - W^*(\mathbf{p}_n)| &\le |W_n - W^*(\hat{\mathbf{p}}_n)| + |W^*(\hat{\mathbf{p}}_n) - W^*(\mathbf{p}_n)| \\ &\le \frac{D}{\sqrt{n}} + \frac{LD\sqrt{M}}{\sqrt{n}}. \end{aligned} \tag{3.24}$$

Finally, since $D = \sqrt{4L^2 + 2}$, we have

$$|W_n - W^*(\mathbf{p}_n)| \le \frac{(L\sqrt{M} + 1)\sqrt{4L^2 + 2}}{\sqrt{n}}. \tag{3.25}$$

Remark. In the proof of the theorem, $W_n \ge W^*(\mathbf{p}_n) - c/\sqrt{n}$, we have identified the associated portfolio

$$\mathbf{b}_{n+1} = \mathbf{b}^*(\hat{\mathbf{p}}_n) \tag{3.26}$$

and the constant

$$c = (L\sqrt{M} + 1)\sqrt{4L^2 + 2}\,, \tag{3.27}$$

where M is the cardinality of $\mathbf{X}$ and $\mathbf{L}$ is the bound on $|\log X_i|$.

4. Example

This algorithm has been tried on the following example. Let $\mathbf{x} \in \mathbf{X} = \{\mathbf{a}_1, \mathbf{a}_2\}$ where

$$\mathbf{a}_1 = (1,2) \quad \text{and} \quad \mathbf{a}_2 = (1,1/2). \tag{4.1}$$

The first component of $\mathbf{x}$ is equal to one, designating cash. An investment in this stock will always be returned. The second component of $\mathbf{x}$ may equal $1/2$ or 2; so the second stock has wild behavior. We first consider log optimal investment with respect to the known distribution

$$P\{\mathbf{X} = \mathbf{a}_1\} = p = 1 - P\{\mathbf{X} = \mathbf{a}_2\}. \tag{4.2}$$

Then,

$$\begin{aligned} W^*(p) &= \max_b \left(p \log \mathbf{b}'\mathbf{a}_1 + (1-p) \log \mathbf{b}'\mathbf{a}_2 \right) \\ &= 0, & p \le 1/3 \\ &= \ln 3 - (1-p)\ln 2 + p \ln p \\ &\qquad + (1-p)\ln(1-p), & 1/3 \le p \le 1/3 \\ &= (2p-1)\ln 2, & p \ge 2/3. \end{aligned} \tag{4.3}$$

The log optimal portfolio is

$$\begin{aligned} \mathbf{b}^* &= (1,0), & p \le 1/3 \\ &= (2-3p, 3p-1), & 1/3 \le p \le 2/3 \\ &= (0,1), & p \ge 2/3. \end{aligned} \tag{4.4}$$

If $p = 1/2$, we notice that neither stock goes anywhere under a buy-and-hold strategy—cash stays at 1, and the factors of 2 and $1/2$ for the wild stock will cancel out (at least to first order in the exponent). Indeed, $W(\mathbf{b}, p) = 0$, for $\mathbf{b} = (1,0)$ and $\mathbf{b} = (0,1)$. Nonetheless, in repeated independent investments, money can be made using $\mathbf{b}^* = (1/2, 1/2)$, yielding

$$W^* = (1/2)\ln(9/8), \tag{4.5}$$

and a growth rate of money

$$S_n = (9/8)^{n/2 + o_p(1)}, \qquad \text{a.e.} \tag{4.6}$$

Money grows at 12.5% every two investment periods.

But what if the drawings of stock vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are not independent and identically distributed? Can one achieve (4.6) if the only available information is $\mathbf{X} = \{\mathbf{a}_1, \mathbf{a}_2\}$ and the past stock outcomes $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, as the theorem asserts? We tried the algorithms $\mathbf{b}^*(\hat{\mathbf{p}}_k)$ on the following four sequences of length 5000:

(1) $\mathbf{a}_1\mathbf{a}_1 \ldots \mathbf{a}_1\mathbf{a}_2 \ldots \mathbf{a}_2$ (2500 $\mathbf{a}_1$'s followed by 2500 $\mathbf{a}_2$'s),

(2) $\mathbf{a}_2\mathbf{a}_2 \ldots \mathbf{a}_2\mathbf{a}_1\mathbf{a}_1 \ldots \mathbf{a}_1$ (2500 $\mathbf{a}_2$'s followed by 2500 $\mathbf{a}_1$'s),

(3) $\mathbf{a}_1\mathbf{a}_2\mathbf{a}_1\mathbf{a}_2 \ldots \mathbf{a}_1\mathbf{a}_2$ (alternating),

(4) $\mathbf{a}_2\mathbf{a}_1\mathbf{a}_2\mathbf{a}_1 \ldots \mathbf{a}_2\mathbf{a}_1$ (alternating).

For each sequence the resulting capital S_{5000} was computed, yielding $(1/5000)\log_2 S_{5000}$ respectively equal to (1) .084864, (2) .084772, (3) .083758, and (4) .083755. Incidentally, this yields $S_{5000} = 2^{418.8}$ units in case 4, a nice improvement over buy-and-hold's $S_{5000} = 1$.

The optimal strategy $\mathbf{b}^* = (1/2, 1/2)$ (with foreknowledge of the proportion of $\mathbf{a}_1$'s) yields

$$\frac{1}{5000} \log_2 S_{5000} = .084962.$$

Observe that none of these sequences fooled the investor into overaggressive behavior.

5. Concluding Remarks

There are obvious hazards in basing the current portfolio on the past behavior $\mathbf{p}_n$ of the market. Certain market sequences may set up the investor for catastrophic future losses. Apparently the empirical distribution $\mathbf{p}_n$ is too sensitive to the past. By replacing $\mathbf{p}_n$ by $\hat{\mathbf{p}}_n$ as developed in the Theorem, one can guarantee that the corresponding log optimal portfolio $\mathbf{b}^*(\hat{\mathbf{p}}_n)$ achieves the desired exponential capital growth $S_n \doteq e^{nW^*(\mathbf{p}_n)}$ uniformly in all market sequences.

References

1. L. Breiman, Optimal gambling systems for favorable games, *in* "Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability," Vol. 1, pp. 65–78, Univ. of California Press, Berkeley, 1961.
2. H. Robbins, Asymptotically subminimax solutions of compound statistical decision problems, *in* "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability," Univ. of California Press, Berkeley, 1951.
3. H. Robbins, An empirical Bayes approach to statistics, *in* "Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability," Vol. 1, pp. 157–164, Univ. of California Press, Berkeley.

4. J. F. HANNAN AND H. ROBBINS, Asymptotic solutions of the compound decision problem, *Ann. Math. Statist.* **26** (1955), 37–51.
5. D. BLACKWELL, Controlled random walks, *J. Proc. Int. Congr. Math.* **3** (1956), 336–338.
6. D. BLACKWELL, An analog of the minimax theorem for vector payoffs, *Pac. J. Math.* **6** (1956), 1–8.
7. P. A. SAMUELSON, General proof that diversification pays, *J. Finan. Quant. Anal.* **2** (March 1967), 1–13.
8. N. HAKANSSON, A characterization of optimal multiperiod portfolio policies, *in* "Portfolio Theory, 25 Years After: Essays in Honor of Harry Markowitz," (Elton and Gruber, Eds.), TIMS Studies in the Management Sciences, Vol. 11, pp. 169–177, North-Holland, Amsterdam/New York/Oxford, 1979.
9. K. ARROW, "Essays in the Theory of Risk Bearing," North-Holland/American Elsevier, 1974.
10. H. MARKOWITZ, Portfolio selection, *J. Finan.* **8**, No. 1 (March 1952), 77–91.
11. J. MOSSIN, Optimal multiperiod portfolio policies, *J. Bus.* **41** (1968), 215–229.
12. W. F. SHARPE, A simplified model for portfolio analysis, *Manage. Sci.* **9** (January 1963), 277–293.
13. E. THORP, Portfolio choice and the Kelly criterion, *Bus. Econ. Statist. Proc. Amer. Stat. Assoc.* (1971), 215–224.
14. R. BELL AND T. M. COVER, Competitive optimality of logarithmic investment, *Math. Oper. Res.* **5**, No. 2 (May 1980), 161–166.
15. T. M. COVER, An algorithm for maximizing expected log investment return, *IEEE Trans. Inf. Theory* **IT-30**, No. 2 (March 1984), 369–373.
16. M. FINKELSTEIN AND R. WHITLEY, Optimal strategies for repeated games, *Adv. Appl. Probab.* **13** (1981), 415–428.